KPR Institute of Engineering and Technology

Learn Beyond

(Autonomous | Affiliated to Anna University)

POSTGRADUATE PROGRAMMES

CURRICULUM AND SYLLABI

PROGRAMME

M.Tech. | Data Sciences

REGULATIONS

2026

CONTROLLED COPY

I. Vision and Mission of the Institute

Vision

To become a premier institute of academic excellence by imparting technical, intellectual and professional skills to students for meeting the diverse needs of industry, society, the nation and the world at large

Mission

- Commitment to offer value-based education and enhancement of practical skills
- Continuous assessment of teaching and learning processes through scholarly activities
- Enriching research and innovation activities in collaboration with industry and institutes of repute
- Ensuring the academic processes to uphold culture, ethics and social responsibilities

II. Vision and Mission of the Department

Vision

To be a department with global recognition through innovative teaching process, product development and research excellence in Artificial Intelligence and Data Science

Mission

The mission of the department is to

- Provide the highest quality of teaching and learning opportunities to the students with a well-researched and carefully crafted curriculum
- Enable the students to develop projects and products in latest technologies related to data science and artificial intelligence with the help of industrial collaboration
- Develop a top notch research culture by empowering students and faculty in fundamentals and applications of data science and artificial intelligence
- Educate the students to apply their skills in solving social challenges in an ethical manner

III. Program Educational Objectives (PEOs)

The Program Educational Objectives (PEOs) of the DATA SCIENCES (DS) represent major accomplishments that the graduates are expected to achieve after two years of graduation.

PEO1: Graduates will apply advanced knowledge and skills in Data Science, Artificial Intelligence, and emerging technologies to analyze complex problems and develop innovative, data-driven solutions for industry and society

PEO2: Graduates will pursue professional growth and research excellence by independently undertaking research, product development, and lifelong learning in Data Science and related multidisciplinary domains

PEO3: Graduates will demonstrate effective communication, leadership, teamwork, and ethical responsibility while addressing societal, environmental, legal, and sustainability challenges through responsible use of data and technology

IV. Program Outcomes (POs)

Graduates of the DATA SCIENCES will be able to

PO1: An ability to independently carry out research /investigation and development work to solve practical problems

PO2: An ability to write and present a substantial technical report/document

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program

PO4: An ability to apply advanced data science techniques, computational tools, and emerging technologies to analyze complex real-world problems and develop effective solutions

PO5: An ability to communicate effectively and work collaboratively with multidisciplinary teams to address technological and societal challenges using data-driven approaches

Head of the Department
Department of Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore - 641 407.

KPR Institute of Engineering and Technology
Centre for Academic Courses
Coimbatore
Learn Beyond

PO6: An ability to recognize and address ethical, societal, legal, and sustainability considerations in the responsible use of data, artificial intelligence, and data science technologies

V. PEO / PO Mapping

Following three levels of correlation should be used:

1: Low

2: Medium

3: High

Particular	PO1	PO2	PO3	PO4	PO5	PO6
PEO 1	3	2	3	3	2	2
PEO 2	3	3	3	3	2	2
PEO 3	2	3	2	2	3	3

VI. Mapping of Course Outcomes (COs) with Program Outcomes (POs)

Sem	Course Title	PO1	PO2	PO3	PO4	PO5	PO6
I	Computational Mathematics for Data Science	✓	✓	✓	-	-	-
	Applied Data Science	✓	✓	✓	✓	✓	✓
	Advanced Machine Learning	✓	✓	✓	✓	✓	✓
	Advanced Data Structures and Algorithms	✓	-	✓	✓	✓	✓
	Research Methodology and IPR	✓	✓	-	-	-	✓
	Advanced Machine Learning Laboratory	✓	✓	✓	✓	✓	✓

Head of the Department
Department of Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore - 641 407.

Learn Beyond

DATA SCIENCES

Regulation 2026

For the students admitted from 2026 onwards

CHOICE BASED CREDIT SYSTEM

CURRICULUM FOR I - IV SEMESTERS

SEMESTER I

No	Course Code	Title	Category	Type	L	T	P	J	C
1	P26MA105	Computational Mathematics for Data Science	FC	TwL	2	-	2	-	3
2	P26DS101	Applied Data Science	PCC	TwL	3	-	2	-	4
3	P26DS102	Advanced Machine Learning	PCC	T	3	-	-	-	3
4	P26CS101	Advanced Data Structures and Algorithms	PCC	TwL	2	-	2	-	3
5	P26RMC01	Research Methodology and IPR	RMC	T	3	-	-	-	3
6	P26DS103	Advanced Machine Learning Laboratory	PCC	L	-	-	4	-	2
7	-	Professional Elective - I	PEC	T	3	-	-	-	3
Total									21

T - Theory, L - Laboratory, TwL - Theory with Laboratory, TwP - Theory with Project

Head of the Department
Department of Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore - 641 407.

Learn Beyond

KPR Institute of Engineering and Technology
Centre for Academic Courses
Coimbatore

PROFESSIONAL ELECTIVE COURSES: VERTICALS

VERTICAL I - DATA SCIENCES

SI.NO.	COURSE CODE	COURSE TITLE	CATEGORY	L	T	P	C
1	P26DSP01	Computer Vision Technologies	PEC	3	0	0	3
2	P26DSP02	Predictive Analytics for IoT	PEC	3	0	0	3
3	P26DSP03	Cyber Security Practices	PEC	3	0	0	3
4	P26DSP04	Applied Accelerated AI	PEC	3	0	0	3
5	P26DSP05	AI TRiSM	PEC	3	0	0	3
6	P26DSP06	NoSQL Database	PEC	3	0	0	3

Head of the Department
Department of Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore - 641 407.

KPR Institute of Engineering and Technology
Centre for Academic Courses
Coimbatore

Learn Beyond

SEMESTER I

		Category: FC				
P26MA105	Computational Mathematics for Data Science	L	T	P	J	C
		2	0	2	0	3

QP PATTERN: B

COURSE OBJECTIVES:

- To build mathematical foundations for data science, covering linear algebra concepts essential for data representation and manipulation
- To develop proficiency in optimization and probability concepts used in machine learning and data analysis
- To apply computational methods in Python for solving mathematical problems in data science

COURSE OUTCOMES:

CO 1:	Explain linear algebra concepts used in data science	Understand
CO 2:	Apply matrix operations and vector space concepts for data representation using Python	Apply
CO 3:	Apply optimization techniques for data-driven problems using Python	Apply
CO 4:	Apply probability concepts in data analysis using Python	Apply
CO 5:	Apply numerical methods to solve computational mathematical problems using Python	Apply

CO - PO MAPPING:

Particular	PO1	PO2	PO3	PO4	PO5	PO6
CO 1	3	2	1	-	-	-
CO 2	3	3	1	-	-	-
CO 3	3	3	2	-	-	-
CO 4	3	3	2	-	-	-
CO 5	3	3	2	-	-	-

SYLLABUS:

UNIT I: LINEAR ALGEBRA 6 + 6

Vectors and Vector Spaces – Linear Independence and Basis – Matrix Operations – Eigenvalues and Eigenvectors – Singular Value Decomposition (SVD) – Dimensionality Reduction Concepts

UNIT II: MULTIVARIATE CALCULUS 6

Partial Derivatives – Gradient and Directional Derivatives – Jacobian Matrix – Hessian Matrix – Taylor Series Expansion for Multivariable Functions

UNIT III: OPTIMIZATION FUNDAMENTALS 6 + 6

Optimization of Multivariable Functions – Lagrange Multiplier – Second Order Optimality Conditions – Positive Definiteness of Hessian

UNIT IV: UNCONSTRAINED OPTIMIZATION TECHNIQUES

6 + 6

Convex Sets and Functions – Gradient Descent Method – Stochastic Gradient Descent – Newton's Method – Conjugate Gradient Method

UNIT V: CONSTRAINED OPTIMIZATION TECHNIQUES

6 + 6

Constrained Optimization – Karush-Kuhn-Tucker (KKT) Conditions – Duality Theory – Primal and Dual Problems

LIST OF EXPERIMENTS

1. Linear Algebra, SVD, and PCA DS Application: Dimensionality Reduction • Perform matrix operations and compute eigenvalues/eigenvectors • Implement SVD on a dataset and analyze singular values • Apply PCA for dimensionality reduction and visualize results
2. Gradient, Hessian, and Taylor Series Computation DS Application: Gradient and Hessian of multivariable functions • Compute gradient and Hessian of multivariable functions using NumPy • Visualize gradient vector field for 2D functions. • Implement Taylor series approximation and verify with analytical results
3. Lagrange Multiplier and Second Order Optimality DS Application: Optimization • Compute gradient of multivariable function using NumPy • Solve constrained optimization using Lagrange multiplier method • Visualize gradient vector field and optimal point
4. Gradient Descent, Newton's Method, and Conjugate Gradient DS Application: Function minimization. • Implement gradient descent for function minimization • Implement Newton's method and conjugate gradient method • Compare convergence rates of all three methods
5. KKT Conditions and Duality Theory DS Application: Inequality constrained problems • Formulate and solve inequality constrained problems • Apply KKT conditions to find optimal solution • Formulate primal and dual problems and verify strong duality

CONTACT PERIODS:

Lecture: 24 Periods **Tutorial:** - Periods **Practical:** 30 Periods **Project:** - Periods **Total:** 54 Periods

TEXTBOOKS:

1. Gilbert Strang, "Introduction to Linear Algebra", 6th Edition, Wellesley-Cambridge Press, 2023
2. James Stewart, "Calculus: Early Transcendentals", 9th Edition, Cengage Learning, 2020
3. Stephen Boyd and Lieven Vandenberghe, "Convex Optimization", Cambridge University Press, 2004
4. S. S. Rao, "Engineering Optimization: Theory and Practice", 5th Edition, John Wiley & Sons, 2019

REFERENCES:

1. David C. Lay, "Linear Algebra and Its Applications", 6th Edition, Pearson, 2021
2. Jerrold E. Marsden and Anthony J. Tromba, "Vector Calculus", 6th Edition, W.H. Freeman, 2011
3. Jorge Nocedal and Stephen J. Wright, "Numerical Optimization", 2nd Edition, Springer, 2006
4. MATLAB Documentation: "Optimization Toolbox", The MathWorks, 2025

Head of the Department
Department of Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology

KPR Institute of Engineering and Technology
Centre for Academic Courses
Coimbatore

SEMESTER I

		Category: PCC				
P26DS101	Applied Data Science (Common to DS, PHDIC)	L	T	P	J	C
		3	0	2	0	4

QP PATTERN: B

COURSE OBJECTIVES:

- To familiarize learners with the tools and techniques needed to gather and organize data from varied sources and file formats using Python
- To build proficiency in preparing, examining and modeling data through preprocessing, statistical analysis and machine learning approaches
- To enable learners to present analytical findings through well-designed visualizations and interactive dashboards

COURSE OUTCOMES:

CO 1:	Acquire, load and store data from diverse sources using appropriate Python libraries and file formats	Apply
CO 2:	Apply data wrangling and preprocessing techniques to clean, transform and reshape raw data	Apply
CO 3:	Analyze data through exploratory data analysis and statistical inference techniques	Analyze
CO 4:	Design, implement and evaluate predictive and machine learning models for practical problems	Evaluate
CO 5:	Develop interactive visualizations and dashboards to communicate data-driven insights effectively	Apply

CO - PO MAPPING:

Particular	PO1	PO2	PO3	PO4	PO5	PO6
CO 1	2	1	2	3	1	2
CO 2	2	1	2	3	1	2
CO 3	3	2	3	3	2	1
CO 4	3	2	3	3	2	2
CO 5	2	3	2	2	3	1

SYLLABUS:

UNIT I: DATA ACQUISITION, LOADING AND STORING 9 + 6

Introduction to Data Science – Python libraries for data processing: NumPy and Pandas – Data Acquisition: gathering information from different sources, Web APIs, Open Data Sources, Web Scraping – Data Loading and Storing: Reading and Writing Data in Text Format, Reading Text Files in Pieces, Writing Data to Text Format, Working with Delimited Formats, JSON Data, XML and HTML

UNIT II: DATA WRANGLING AND PREPROCESSING **9 + 6**

Data Wrangling: Clean, Transform, Merge, Reshape – Combining and Merging Datasets, Merging on Index, Concatenation, Combining with Overlap, Reshaping and Pivoting, Group-by operations on DataFrames – Pivot Tables and Crosstabs for understanding relationships between variables – Handling Missing Data, Data Transformation, String Manipulation, Summarizing, Binning, Classing and Standardization – Outliers, Noise and Anomalies, Encoding Techniques and Choice of Encoding

UNIT III: EXPLORATORY DATA ANALYSIS AND STATISTICAL INFERENCE **9 + 6**

Exploratory Data Analysis – Probability and Statistics: Rectangular Data – Estimates of Location – Estimates of Variability, Exploring Data Distribution – Exploring Binary and Categorical Data, Correlation, Exploring Two or More Variables – Time Series Analysis – Random Sampling and Sample Bias, Sampling Distribution of a Statistic, The Bootstrap, Confidence Intervals – Statistical Experiments and Significance Testing: A/B Testing, Hypothesis Tests, Resampling, Statistical Significance and p-Values

UNIT IV: PREDICTIVE MODELING AND MACHINE LEARNING **9 + 6**

Modeling and Prediction: Regression and Prediction, Classification techniques – Machine Learning models: K-Nearest Neighbors – Tree Models – Bagging and Random Forest, Boosting – Unsupervised Learning: concepts and applications – Model selection and performance considerations for real-world datasets

UNIT V: DATA VISUALIZATION AND STORYTELLING **9 + 6**

Visualization: Data Visualization Principles, Power BI and Tableau Basics, Dashboard Design Principles, Interactive Visualization Techniques – Storytelling with Data – Case Studies in effective data communication

LIST OF EXPERIMENTS

1. Data Processing Operations using NumPy and Pandas – UrbanMart Retail Sales Analysis. UrbanMart records daily store-wise revenue across a week. Use NumPy for array-based computation and Pandas for DataFrame-based analysis to identify underperforming stores and the best sales day
2. Data Acquisition from Web APIs and Open Data Sources – CodeVault Package Release Tracker. CodeVault wants to track the release history of a key open-source dependency to plan its upgrade cycles. Acquire version and release-date data from the free, key-less PyPI JSON API and load them into a Pandas DataFrame.
3. Web Scraping using BeautifulSoup – DevScout Trending Repository Tracker. DevScout tracks trending open-source repositories for a given technology topic. Scrape repository owner, name, description and star count from a GitHub Topics page using requests and BeautifulSoup, then summarise the listing with Pandas
4. Programs using Regular Expression Functions and Classes – Telecom Support Ticket Parser. A telecom support team forwards raw ticket text. Build a TicketParser class using Python's re module to extract email addresses, phone numbers and ticket reference codes from each ticket
5. Read and Write Data Across File Formats – Library Catalogue Migration. A library's book records are scattered across Text, CSV, JSON, XML and HTML sources. Read each format, map it to a common schema, and merge all five sources into one unified catalogue
6. Group-by, Pivot Table and Crosstab based Aggregation – HR Appraisal Analytics. Before the annual appraisal cycle, HR wants workforce and rating summaries by department, designation and gender using groupby, pivot_table and crosstab

7. Data Transformation, Summarizing and Encoding Techniques for Categorical Data – Bank Loan Approval Dataset. A bank's loan-approval dataset mixes numeric and categorical fields. Summarise the categorical columns, then apply label encoding to the ordinal Education field and one-hot encoding to the nominal fields
8. Binning, Classing, Standardization and Outlier / Anomaly Handling – FitPulse Sensor Data. FitPulse collects raw heart-rate, step-count and calorie data. Bin heart rate into categories, standardise steps and calories, and detect outliers using both the IQR and z-score methods
9. Exploratory Data Analysis – Estimates of Location, Variability and Data Distribution – Student Marks. A university's academic office wants to understand semester marks before deciding on remedial classes. Compute location and variability estimates and inspect each subject's distribution shape
10. Statistical Significance Testing – A/B Testing and Confidence Intervals – Checkout Redesign. An e-commerce product team ran a redesigned checkout (Version B) against the original (Version A). Test whether the difference in conversion rate is statistically significant and build a confidence interval for it
11. Implementation and Evaluation of Predictive Models – Bank Loan Regression, Classification and Ensemble Methods. A bank wants to predict the loan amount to sanction and whether an applicant will default. Build a regression model, two classification models, and an ensemble, then compare their performance
12. Interactive Dashboard Design and Data Storytelling using Power BI – Executive Sales Dashboard. The CEO wants a single interactive dashboard reviewed every Monday, summarising last week's sales performance across regions and categories, with a short narrative on what changed

CONTACT PERIODS:

Lecture: 45 Periods **Tutorial:** - Periods **Practical:** 30 Periods **Project:** - Periods **Total:** 75 Periods

TEXTBOOKS:

1. Peter Bruce, Andrew Bruce, and Peter Gedeck, "Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python", O'Reilly Media
2. Wes McKinney, "Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Jupyter", 3rd Edition, O'Reilly Media, 2022, ISBN-13: 978-1098104030
3. Alex Kolokolov and Maxim Zelensky, "Data Visualization with Microsoft Power BI: How to Design Savvy Dashboards", 1st Edition, O'Reilly Media, 2024, ISBN-13: 978-1098152789

REFERENCES:

1. Joel Grus, "Data Science from Scratch: First Principles with Python", 2nd Edition, O'Reilly Media, 2019
2. Jake VanderPlas, "Python Data Science Handbook: Essential Tools for Working with Data", O'Reilly Media, 2016
3. Cathy O'Neil and Rachel Schutt, "Doing Data Science", O'Reilly Media, 2013
4. Ani Adhikari and John DeNero, "Computational and Inferential Thinking: The Foundations of Data Science", e-book

Head of the Department
Department of Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore - 641 407.

KPR Institute of Engineering and Technology
Centre for Academic Courses
Coimbatore
Learn Beyond

SEMESTER I

P26DS102	Advanced Machine Learning	Category: PCC				
		L	T	P	J	C
		3	0	0	0	3

QP PATTERN: C

COURSE OBJECTIVES:

- To develop a strong theoretical foundation in machine learning principles, model evaluation techniques, and probabilistic learning methods to prepare students for real-world data-driven problem solving
- To equip students with hands-on skills to design, implement, and evaluate supervised and unsupervised learning models using industry-standard algorithms and performance metrics
- To enable students to engineer high-performance machine learning pipelines by applying advanced techniques in feature engineering, neural networks, and ensemble learning for software and engineering applications

COURSE OUTCOMES:

CO 1: Demonstrate a thorough understanding of machine learning types, hypothesis space, and bias–variance trade-off to evaluate and select appropriate models for a given problem — **Apply**

CO 2: Analyse supervised machine learning algorithms to determine appropriate models for regression and classification problems — **Analyze**

CO 3: Analyse unsupervised learning techniques to select appropriate clustering methods for diverse datasets — **Analyze**

CO 4: Use dimensionality reduction and feature selection methods including PCA, LDA, and ICA to improve the efficiency and accuracy of machine learning pipelines — **Apply**

CO 5: Build and deploy advanced predictive systems using neural networks and ensemble methods such as bagging and boosting to solve real-world engineering problems — **Apply**

CO - PO MAPPING:

Particular	PO1	PO2	PO3	PO4	PO5	PO6
CO 1	3	2	1	3	2	1
CO 2	3	2	1	3	2	1
CO 3	3	2	1	3	2	1
CO 4	3	2	1	3	2	1
CO 5	3	2	1	3	2	1

SYLLABUS:

UNIT I: MACHINE LEARNING - FOUNDATION — **9**

Introduction to ML – Types – Types of learning: supervised, unsupervised, semi-supervised, and reinforcement learning – Hypothesis space and inductive bias – Model evaluation: train/test split, cross-validation, and performance metrics – Bias–variance trade-off – Generalisation versus overfitting

Learn Beyond

UNIT II: SUPERVISED LEARNING 9

Regression: Linear regression, multiple linear regression, and polynomial regression – Classification: logistic regression and maximum likelihood estimation – Decision trees: CART for both classification and regression tasks – Support vector machines: linear and non-linear kernel functions – K-nearest neighbours (KNN) – Performance measures for classification and prediction models

UNIT III: UNSUPERVISED AND PROBABILISTIC LEARNING 9

Clustering: Introduction to clustering – Partition-based clustering (K-Means) – Hierarchical clustering (Agglomerative) – Density-based clustering (DBSCAN) – Distance and similarity measures – Introduction to Bayesian learning – Naïve Bayes classifier – Applications of clustering and Naïve Bayes

UNIT IV: FEATURE ENGINEERING AND DIMENSIONALITY REDUCTION 9

Feature Engineering: Introduction to feature selection and feature extraction – Feature selection methods – Feature extraction techniques – Principal Component Analysis (PCA) for dimensionality reduction – Introduction to Linear Discriminant Analysis (LDA) – Applications of feature engineering in machine learning

UNIT V: NEURAL NETWORKS AND ENSEMBLE LEARNING 9

Artificial neural networks: Perceptron and multi-layer perceptron (MLP) – Backpropagation and training deep networks – Applications of neural networks in software engineering – Ensemble learning: bagging (Random Forest), boosting (AdaBoost, gradient boosting) – Applications of ensemble methods in software engineering contexts

CONTACT PERIODS:

Lecture: 45 Periods **Tutorial:** - Periods **Practical:** - Periods **Project:** - Periods **Total:** 45 Periods

TEXTBOOKS:

1. Murphy, K. P, "Probabilistic machine learning: An introduction", MIT Press, 2022
2. Mitchell, T. M, "Machine learning", Vol. 1 , McGraw-hill New York, 1997

REFERENCES:

1. Christopher M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006
2. Trevor Hastie, Robert Tibshirani, and Jerome Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd edition, Springer, 2009
3. Richard S. Sutton and Andrew G. Barto, Reinforcement Learning: An Introduction, 2nd edition, MIT Press, 2018

Head of the Department
Department of Artificial Intelligence and Data Science
Coimbatore - 641 407.

KPR Institute of Engineering and Technology
Centre for Academic Courses
Coimbatore

SEMESTER I

P26CS101	Advanced Data Structures and Algorithms (Common to DS, PCS, PHDIC)	Category: PCC				
		L	T	P	J	C
		2	0	2	0	3

QP PATTERN: B

COURSE OBJECTIVES:

- To acquire knowledge in role of algorithms and tree structures
- To gain knowledge on the graphs and linear programming
- To understand the concept of NP completeness and approximation algorithms

COURSE OUTCOMES:

CO 1:	Determine the time complexity of algorithms using asymptotic notations and mathematical methods	**Apply**
CO 2:	Perform efficient insertion, deletion, and searching operations using binary search trees, red-black trees, B-trees, and Fibonacci heaps	**Apply**
CO 3:	Solve graph-based problems using graph traversal, minimum spanning tree, and single-source shortest path algorithms	**Apply**
CO 4:	Evaluate the performance of multithreaded algorithms and formulate solutions using linear programming techniques for complex computational problems	**Analyze**
CO 5:	Classify computational problems based on NP-completeness and formulate near-optimal solutions using approximation algorithms	**Analyze**

CO - PO MAPPING:

Particular	PO1	PO2	PO3	PO4	PO5	PO6
CO 1	2	-	3	-	3	2
CO 2	2	-	3	-	3	2
CO 3	2	-	3	1	3	2
CO 4	2	-	3	-	3	2
CO 5	2	-	3	1	3	2

SYLLABUS:

UNIT I: ROLE OF ALGORITHMS IN COMPUTING **6 + 6**

Algorithms – Algorithms as a Technology – Insertion Sort – Analyzing Algorithms – Designing Algorithms – Growth of Functions: Asymptotic Notation – Standard Notations and Common Functions-Recurrences: The Substitution Method – The Recursion-Tree Method

UNIT II: HIERARCHICAL DATA STRUCTURES **6 + 6**

Binary Search Trees: Basics – Querying a Binary search tree – Insertion and Deletion- Red-Black trees: Properties of Red-Black Trees – Rotations – Insertion – Deletion -B-Trees: Definition of Btrees – Basic operations on B-Trees – Deleting a key from a B-Tree – Fibonacci Heaps: structure – Mergeable-heap operations

UNIT III: GRAPHS **6 + 6**

Elementary Graph Algorithms: Representations of Graphs – Breadth-First Search – Depth-First Search – Topological Sort – Strongly Connected Components- Minimum Spanning Trees: Growing a Minimum Spanning Tree – Kruskal and Prim- Single-Source Shortest Paths: The Bellman-Ford algorithm – Single-Source Shortest Paths in Directed Acyclic Graphs – Dijkstra's Algorithm

UNIT IV: MULTITHREADED ALGORITHMS AND LINEAR PROGRAMMING **6 + 6**

Dynamic multithreading – Multithreaded matrix multiplication – Multithreaded merge sort – Linear Programming – Standard and slack forms – Formulating problems as linear programs – Simplex algorithm – Duality – Feasible solution

UNIT V: NP-COMPLETENESS AND APPROXIMATION ALGORITHMS **6 + 6**

Polynomial time – verification – NP-completeness and reducibility – NP-completeness proofs – NP complete problems – Approximation Algorithms – Vertex-cover problem – Traveling-salesman problem – Set-covering problem – Randomization and linear programming – Subset-sum problem

LIST OF EXPERIMENTS

1. A company has an unsorted list of employee records (ID, salary). Use Insertion Sort to arrange them in ascending order of salary and analyze how sorting time grows as the number of employees increases
2. Design a Binary Search Tree–based contact book that stores names/phone numbers, allowing fast insertion of new contacts, deletion of old ones, and quick lookup — and compare search time against a simple unsorted list
3. Simulate how a database index (e.g., for a banking application with continuous insertions of transaction records) stays balanced using a Red-Black Tree, so search/insert operations remain efficient even under heavy, frequent updates
4. Simulate how an operating system or DBMS indexes large files on disk using a B-Tree, where each node represents a disk block — demonstrating why B-Trees are preferred over BSTs for large-scale disk-based storage
5. Model a social network (or website links) as a graph and use BFS to suggest "friends of friends," and DFS to detect connected communities or crawl through the network, including topological sorting for task/course-prerequisite scheduling
6. Given a set of cities/villages and the cost of laying cables/wires between them, use Kruskal's and Prim's algorithms to design the minimum-cost network connecting all locations
7. Model a city's road network as a weighted graph and use Dijkstra's algorithm to find the shortest route between two locations (like Google Maps); use Bellman-Ford to handle a network with toll discounts/negative edge weights or detect fare-scam negative cycles
8. Use multithreaded matrix multiplication to speed up a real workload such as image transformation or scientific computation, and multithreaded merge sort to sort large-scale sensor/log data faster using multi-core processors
9. A factory wants to maximize profit given limited raw materials, labor hours, and machine time. Formulate this as a Linear Program (standard/slack form) and solve using the Simplex algorithm to find the optimal product mix
10. Use an approximation algorithm for the Vertex-Cover problem to place the minimum number of CCTV cameras/security guards covering all road junctions in a city
11. Use an approximation algorithm for the Traveling Salesman Problem to plan the shortest delivery route for a courier/ logistics company visiting multiple cities

CONTACT PERIODS:

Lecture: 30 Periods **Tutorial:** - Periods **Practical:** 30 Periods **Project:** - Periods **Total:** 60 Periods

TEXTBOOKS:

1. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest and Clifford Stein, Introduction to Algorithms, 4th Edition, The MIT Press, 2022
2. Jon Kleinberg and Éva Tardos, Algorithm Design, 1st Edition, Pearson, 2005

REFERENCES:

1. Robert Sedgewick and Kevin Wayne, Algorithms, 4th Edition, Addison-Wesley Professional, 2011. The book covers algorithm analysis, searching, balanced search trees, graph algorithms, minimum spanning trees and shortest paths
2. Sanjoy Dasgupta, Christos H. Papadimitriou and Umesh Vazirani, Algorithms, 1st Edition, McGraw-Hill, 2006

Head of the Department
Department of Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore - 641 407.

KPR Institute of Engineering and Technology
Centre for Academic Courses
Coimbatore

SEMESTER I

P26RMC01	Research Methodology and IPR (Common to CAD, DS, ES, PCS, PHDCE, PHDIC, PHDME, PHDSH, SE, VL)	Category: RMC				
		L	T	P	J	C
		3	0	0	0	3

QP PATTERN: B

COURSE OBJECTIVES:

- To impart knowledge in problem formulation, analysis and solutions
- To impart skills required for technical paper writing / presentation without violating professional ethics
- To familiarize knowledge on patent drafting and filing patents

COURSE OUTCOMES:

CO 1:	Apply the principles of research problem formulation to identify, define, and investigate engineering research problems using appropriate data collection and analysis techniques	Apply
CO 2:	Analyze research literature, data collection methods, research ethics, and plagiarism issues to identify research gaps and select suitable methodologies	Analyze
CO 3:	Evaluate the quality of technical reports, research papers, and research proposals by applying standard scientific writing and presentation practice	Evaluate
CO 4:	Apply the principles and procedures of Intellectual Property Rights (IPR), including patents, patent filing, and international patenting systems, to protect technological innovation	Apply
CO 5:	Evaluate patent rights, technology transfer mechanisms, patent databases, geographical indications, and emerging developments in Intellectual Property Rights for engineering and research applications	Evaluate

CO - PO MAPPING:

Particular	PO1	PO2	PO3	PO4	PO5	PO6
CO 1	3	-	-	-	-	-
CO 2	3	-	-	-	-	-
CO 3	3	3	-	-	-	-
CO 4	-	3	-	-	-	2
CO 5	-	3	-	-	-	2

SYLLABUS:

UNIT I: RESEARCH PROBLEM FORMULATION 9

Meaning of research problem – Sources of research problem – Criteria characteristics of a good research problem – Errors in selecting a research problem – Scope and objectives of research problem. Approaches of investigation of solutions for research problem – Data collection – Analysis – Interpretation – Necessary instrumentations

UNIT II: LITERATURE REVIEW AND DATA COLLECTION 9

Effective literature studies approaches – Analysis – Plagiarism and research ethics. Method of data collection – Types of data – Primary Data – Scales of measurement – Source and collection of data observation method – Secondary data

UNIT III: TECHNICAL WRITING / PRESENTATION 9

Effective technical writing: How to write report – Paper – Developing a research proposal – Format of research proposal – Presentation and assessment by a review committee

UNIT IV: INTRODUCTION TO INTELLECTUALPROPERTY RIGHTS 9

Nature of Intellectual Property: Patents – Designs – Trade and Copyright. Process of Patenting and Development – technological research – Innovation – Patenting – Development – International Scenario – International cooperation on Intellectual Property – Procedure for grants of patents – Patenting under PCT

UNIT V: INTELLECTUAL PROPERTYRIGHTS 9

Patent Rights: Scope of Patent Rights – Licensing and transfer of technology – Patent information and databases – Geographical Indications – New Developments in IPR – Administration of Patent System – IPR of Biological Systems – Computer Software etc.

CONTACT PERIODS:

Lecture: 45 Periods **Tutorial:** - Periods **Practical:** - Periods **Project:** - Periods **Total:** 45 Periods

REFERENCES:

1. Cooper, DR and Schindler, P S., "Business Research Methods", 9th edition, Tata McGraw Hill, 2014
2. Robert P.Merges, Peter S, Menell, Mark A.Lemley, "Intellectual Property in New Technological age", Volume I, Clause 8 Publishing,2022
3. Ranjit Kumar, "Research Methodology: A Step by Step Guide for beginners", 3rd edition, Sage Publications,2010

Head of the Department
Department of Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore - 641 407.

KPR Institute of Engineering and Technology
Centre for Academic Courses
Coimbatore

Learn Beyond

SEMESTER I

		Category: PCC				
P26DS103	Advanced Machine Learning Laboratory	L	T	P	J	C
		0	0	4	0	2

QP PATTERN: B

COURSE OBJECTIVES:

- To impart practical skills to students to implement and evaluate supervised and unsupervised learning models on real-world datasets
- To train students in applying feature engineering and dimensionality reduction techniques to build efficient and interpretable machine learning pipelines
- To enable students to design and deploy advanced machine learning systems using neural networks and ensemble methods for solving complex software and data engineering problems

COURSE OUTCOMES:

CO 1:	To implement fundamental machine learning workflows including data splitting, cross-validation, and bias–variance analysis using real-world datasets to assess model generalisation	Apply
CO 2:	To develop supervised learning models such as logistic regression, decision trees, support vector machines, and KNN for solving classification and prediction problems on publicly available datasets	Apply
CO 3:	To build unsupervised learning pipelines using clustering algorithms and Naïve Bayes classifiers to discover hidden patterns and perform probabilistic inference on unlabelled real-world data	Apply
CO 4:	To apply dimensionality reduction and feature selection techniques including PCA and LDA on high-dimensional datasets to improve model interpretability and predictive performance	Apply
CO 5:	To design and evaluate advanced machine learning systems using multi-layer perceptrons and ensemble methods such as Random Forest and AdaBoost for defect prediction and image classification tasks	Apply

CO - PO MAPPING:

Particular	PO1	PO2	PO3	PO4	PO5	PO6
CO 1	3	2	3	3	2	2
CO 2	3	2	3	3	2	2
CO 3	3	2	3	3	2	2
CO 4	3	2	3	3	2	2
CO 5	3	2	3	3	2	2

SYLLABUS:

LIST OF EXPERIMENTS

1. Exploring Bias-Variance Trade-off on Housing Price Data Dataset: Boston Housing / Ames Housing (Kaggle) Task: Fit polynomial regression models of increasing degree on housing price data. Plot training vs. validation error curves to visually demonstrate underfitting and overfitting. Evaluate using k-fold cross-validation and compare bias-variance behaviour across model complexities

2. Predicting Loan Default using Logistic Regression Dataset: Home Credit Default Risk (Kaggle) Task: Build a binary classifier to predict whether a loan applicant will default. Apply logistic regression, tune the decision threshold, and evaluate using confusion matrix, precision, recall, F1-score, and ROC-AUC curve
3. Customer Churn Prediction using Decision Trees and KNN Dataset: Telco Customer Churn (Kaggle) Task: Predict telecom customer churn using a CART decision tree and K-nearest neighbours. Compare both models on accuracy, visualise the decision tree structure, and analyse which features drive churn most strongly
4. Cancer Tumour Classification using Support Vector Machines Dataset: Wisconsin Breast Cancer Dataset (UCI / Kaggle) Task: Classify tumours as malignant or benign using SVMs with linear and RBF (Gaussian) kernels. Compare kernel performance, tune hyperparameters using grid search, and evaluate using cross-validation
5. Customer Segmentation using Clustering Algorithms Dataset: Mall Customer Segmentation / E-Commerce Data (Kaggle) Task: Apply k-means, agglomerative hierarchical clustering, and DBSCAN to segment customers based on spending behaviour and demographics. Use elbow method and silhouette score to select optimal k. Visualise clusters using 2D plots
6. Email Spam Detection using Naïve Bayes Dataset: SMS Spam Collection Dataset (Kaggle / UCI) Task: Build a text classification pipeline to detect spam emails. Apply Multinomial Naïve Bayes on TF-IDF features, evaluate with precision and recall, and explore how prior probabilities affect classification outcomes
7. Face Recognition Dimensionality Reduction using PCA and LDA Dataset: Labeled Faces in the Wild / Olivetti Faces (scikit-learn) Task: Apply PCA to reduce high-dimensional facial image data and reconstruct faces at varying numbers of components. Compare PCA with LDA for classification separability. Visualise explained variance and eigenfaces
8. Credit Card Fraud Detection with Feature Selection Dataset: Credit Card Fraud Detection (Kaggle) Task: Handle a highly imbalanced dataset using SMOTE or undersampling. Apply filter-based (chi-square, correlation), wrapper-based (RFE), and embedded (Lasso) feature selection methods. Compare model performance before and after feature selection using logistic regression or random forest
9. Image Classification using Multi-Layer Perceptron Dataset: MNIST Handwritten Digits / Fashion-MNIST (Kaggle / TensorFlow datasets) Task: Build and train an MLP from scratch using backpropagation to classify handwritten digits or clothing items. Experiment with hidden layer sizes, activation functions, learning rates, and dropout. Plot training/validation loss and accuracy curves
10. Software Bug Prediction using Ensemble Learning Dataset: NASA Software Defect Datasets — CM1, KC1, PC1 (Promise Repository / Kaggle) Task: Predict whether a software module contains defects using Random Forest (bagging) and AdaBoost / Gradient Boosting (boosting). Compare ensemble methods against a single decision tree baseline. Evaluate using precision, recall, and AUC. Analyse feature importance to identify which code metrics most predict bugs

CONTACT PERIODS:

Lecture: - Periods **Tutorial:** - Periods **Practical:** 4 Periods **Project:** - Periods **Total:** 4 Periods

TEXTBOOKS:

1. Murphy, K. P, "Probabilistic machine learning: An introduction", MIT Press, 2022
2. Mitchell, T. M, "Machine learning", Vol. 1 , McGraw-hill New York, 1997

REFERENCES:

1. Christopher M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006
2. Trevor Hastie, Robert Tibshirani, and Jerome Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd edition, Springer, 2009
3. Richard S. Sutton and Andrew G. Barto, Reinforcement Learning: An Introduction, 2nd edition, MIT Press, 2018

Head of the Department
Department of Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore - 641 407

KPR Institute of Engineering and Technology
Centre for Academic Courses
Coimbatore

VERTICAL: Data Sciences

		Category: PEC				
P26DSP01	Computer Vision Technologies	L	T	P	J	C
		3	0	0	0	3

QP PATTERN: B

COURSE OBJECTIVES:

- To understand fundamental concepts of image formation, processing, and computer vision pipelines
- To apply feature extraction, motion analysis, and recognition techniques for real-world problems
- To analyse advanced vision systems including 3D reconstruction and deep learning-based approaches

COURSE OUTCOMES:

CO 1:	Apply geometric and photometric image formation models, image filtering, edge detection, transformations, and image pyramids to preprocess and enhance images for basic computer vision applications	Apply
CO 2:	Apply feature detection and model fitting techniques to solve vision problems	Apply
CO 3:	Apply deep learning techniques for image classification and object recognition	Apply
CO 4:	Analyse motion estimation and image alignment in dynamic vision systems	Analyze
CO 5:	Analyse 3D vision techniques and reconstruction methods for real-world applications	Analyze

CO - PO MAPPING:

Particular	PO1	PO2	PO3	PO4	PO5	PO6
CO 1	1	-	2	3	-	-
CO 2	2	-	2	3	1	1
CO 3	2	-	3	3	1	-
CO 4	2	-	3	3	1	2
CO 5	3	1	3	3	1	2

SYLLABUS:

UNIT I: IMAGE FORMATION AND PROCESSING 9

Introduction to Computer Vision – Image Formation: Geometric transformations – Camera models – Photometric image formation – Image Processing – Filtering – Edge detection – Image transformations – Pyramids

UNIT II: FEATURE DETECTION AND MODEL FITTING 9

Model Fitting and Optimization - Least squares – RANSAC – Optimization techniques – Feature Detection and Matching – Edge detection – Corner detection – SIFT/SURF – Feature matching

UNIT III: IMAGE RECOGNITION 9

Neural Networks Fundamentals – Deep Learning for Vision – Convolutional Neural Networks (CNNs) – Training Strategies and Optimization – Image Classification – Object Recognition

UNIT IV: MOTION AND IMAGE ALIGNMENT 9

Image Alignment and Stitching – Feature based alignment – Panorama creation – Motion Estimation – Optical flow – Tracking – Motion models – Computational Photography – HDR imaging – Image enhancement techniques

UNIT V: 3D VISION AND RECONSTRUCTION 9

Stereo Vision – Depth Estimation – Depth Sensing – 3D Reconstruction – Surface Reconstruction – Structure from Motion (SfM) and SLAM – Vision Transformers (ViTs) – Self-Supervised Learning in Vision – Generative AI in Vision (GANs, Diffusion Models)

CONTACT PERIODS:

Lecture: 45 Periods **Tutorial:** - Periods **Practical:** - Periods **Project:** - Periods **Total:** 45 Periods

TEXTBOOKS:

1. Richard Szeliski, "Computer Vision: Algorithms and Applications", Springer, London, 2nd Edition, 2022

REFERENCES:

1. Simon J. D. Prince, "Computer Vision: Models, Learning, and Inference", Cambridge University Press, 2012

2. David A. Forsyth and Jean Ponce, "Computer Vision: A Modern Approach", 2nd Edition, Pearson, 2011

3. E. R. Davies, "Computer Vision: Principles, Algorithms and Practicalities", 4th Edition, Academic Press, 2012

Head of the Department
Department of Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology

KPR Institute of Engineering and Technology
Centre for Academic Courses
Coimbatore

Learn Beyond

VERTICAL: Data Sciences

		Category: PEC				
P26DSP02	Predictive Analytics for IoT	L	T	P	J	C
		3	0	0	0	3

QP PATTERN: B

COURSE OBJECTIVES:

- To understand IoT architecture and data generation
- To learn data preprocessing and analytics techniques
- To apply predictive models for IoT data
- To analyze patterns and knowledge discovery
- To explore IoT applications using analytics

COURSE OUTCOMES:

CO 1:	Explain IoT architecture, components, and data characteristics	**Understand**
CO 2:	Apply data pre-processing and data mining techniques on IoT data	**Apply**
CO 3:	Build and evaluate predictive models using classification and regression techniques	**Analyze**
CO 4:	Analyze patterns using association rule mining and knowledge discovery methods	**Analyze**
CO 5:	Evaluate IoT analytics applications and address data management and security issues	**Evaluate**

CO - PO MAPPING:

Particular	PO1	PO2	PO3	PO4	PO5	PO6
CO 1	-	-	2	2	1	-
CO 2	2	-	3	3	1	-
CO 3	3	-	3	3	1	-
CO 4	3	-	3	3	1	-
CO 5	2	1	3	3	2	3

SYLLABUS:

UNIT I: INTRODUCTION TO IOT 9

Introduction to Internet of Things – IoT architecture and components – Sensors and data acquisition – IoT communication models – Characteristics of IoT data – IoT applications – Challenges in IoT systems – Overview of IoT platforms – Case studies

UNIT II: DATA PREPROCESSING AND DATA MINING CONCEPTS 9

Data collection and storage – Data preprocessing: cleaning, integration, transformation – Data reduction techniques – Data warehouse concepts – Introduction to data mining – Types of data mining tasks – Data mining process – Applications of data mining

UNIT III: PREDICTIVE MODELING TECHNIQUES 9

Classification: Decision Trees, Bayesian classification – Regression analysis – K-Nearest Neighbor method – Model evaluation metrics (accuracy, precision, recall) – Overfitting and model validation – Introduction to clustering (basic idea)

UNIT IV: PATTERN MINING AND KNOWLEDGE DISCOVERY 9

Association rule mining – Frequent itemset mining – Apriori algorithm – Pattern evaluation – Concept of knowledge discovery in databases (KDD) – Data mining applications in IoT – Case studies

UNIT V: IOT ANALYTICS AND APPLICATIONS 9

IoT data analytics overview – Cloud-based IoT analytics – Data management in IoT – Smart city applications – Smart healthcare systems – Industrial IoT – Security and privacy issues in IoT – Future trends in IoT

CONTACT PERIODS:

Lecture: 45 Periods **Tutorial:** - Periods **Practical:** - Periods **Project:** - Periods **Total:** 45 Periods

TEXTBOOKS:

1. Han, J., Kamber, M., Pei, J., "Data Mining: Concepts and Techniques", 4th Edition, Morgan Kaufmann, 2022
2. Buyya, R., Dastjerdi, A. V., "Internet of Things: Principles and Paradigms", 1st Edition, Morgan Kaufmann, 2016

REFERENCES:

1. Tan, P.-N., Steinbach, M., Kumar, V., "Introduction to Data Mining", 2nd Edition, Pearson, 2019
2. McKinney, W., "Python for Data Analysis", 2nd Edition, O'Reilly, 2017
3. Aggarwal, C. C., "Data Mining: The Textbook", 1st Edition, Springer, 2015

Head of the Department
Department of Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore - 641 407.

KPR Institute of Engineering and Technology
Centre for Academic Courses
Coimbatore

Learn Beyond

VERTICAL: Data Sciences

		Category: PEC				
P26DSP03	Cyber Security Practices	L	T	P	J	C
		3	0	0	0	3

QP PATTERN: B

COURSE OBJECTIVES:

- To introduce fundamental concepts of cyber security and the evolving threat landscape
- To understand various cyber attacks, vulnerabilities, and defense mechanisms
- • To explore cryptography, network security practices, and secure system design

COURSE OUTCOMES:

CO 1:	Analyse the fundamental concepts and importance of cyber security in real-world contexts	Analyze
CO 2:	Analyse common cyber threats, vulnerabilities, and attack techniques to assess their impact	Analyze
CO 3:	Apply cryptographic techniques to secure data and communication	Apply
CO 4:	Apply network and system security practices to protect digital asset	Apply
CO 5:	Analyse legal, ethical, and privacy issues in cyber security to evaluate compliance and risk	Analyze

CO - PO MAPPING:

Particular	PO1	PO2	PO3	PO4	PO5	PO6
CO 1	2	-	3	1	-	1
CO 2	2	1	3	2	-	1
CO 3	2	-	3	3	-	-
CO 4	2	1	3	3	1	1
CO 5	1	2	2	1	2	3

SYLLABUS:

UNIT I: INTRODUCTION TO CYBER SECURITY 9

Basics of Cyber Security – CIA Triad (Confidentiality, Integrity, Availability) – Cyber threats, vulnerabilities, and risks – Types of attackers (hackers, insiders, script kiddies) – Cyber attack lifecycle

UNIT II: CYBER ATTACKS AND VULNERABILITIES 9

Malware: viruses, worms, trojans, ransomware – Phishing and social engineering attacks – Password attacks and authentication threats-Web application attacks (SQL Injection, XSS) – Vulnerability assessment and penetration testing basics-Security tool: Wireshark,: Burp Suite, SQLMap, Latest Security Threats: AI-powered cyberattacks, supply chain attacks, deepfake-based social engineering, zero-day exploits, credential stuffing

UNIT III: CRYPTOGRAPHY AND DATA SECURITY 9

Fundamentals of cryptography – Symmetric and asymmetric encryption – Hashing and digital signatures Public Key Infrastructure (PKI) – Secure communication protocols (SSL/TLS basics)

UNIT IV: NETWORK AND SYSTEM SECURITY PRACTICES 9

Network security concepts (firewalls, IDS/IPS) – Secure network design and architecture – Operating system security – Cloud security basics – Security tools and best practices: DoS/DDoS Mitigation – WAF Web Application Firewall.

UNIT V: CYBER LAWS, ETHICS AND RISK MANAGEMENT 9

Cyber laws and regulations (IT Act, GDPR overview) – Ethical hacking and responsible disclosure – Privacy and data protection – Risk management and security policies – Incident response and disaster recovery

CONTACT PERIODS:

Lecture: 45 Periods **Tutorial:** - Periods **Practical:** - Periods **Project:** - Periods **Total:** 45 Periods

TEXTBOOKS:

1. Charles J. Brooks, Christopher Grow, Philip Craig, Donald Short, "Cyber Security Essentials" Wiley, 2018
2. William Stallings "Cryptography and Network Security: Principles and Practice", 7th Edition, Pearson, 2017

REFERENCES:

1. William Stallings and Lawrie Brown, "Computer Security: Principles and Practice", Pearson, 2018
2. Dafydd Stuttard and Marcus Pinto, "The Web Application Hacker's Handbook", Wiley, 2011
3. Jon Erickson, "Hacking: The Art of Exploitation", No Starch Press, 2008
4. Ross Anderson, "Security Engineering", Wiley, 2020

Head of the Department
Department of Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore - 641 407

KPR Institute of Engineering and Technology
Centre for Academic Courses
Coimbatore

VERTICAL: Data Sciences

		Category: PEC				
P26DSP04	Applied Accelerated AI	L	T	P	J	C
		3	0	0	0	3

QP PATTERN: B

COURSE OBJECTIVES:

- To develop in-depth understanding of accelerated computing architectures and AI system design
- To analyze and optimize AI workloads using parallel and distributed computing techniques
- To evaluate scalable AI deployment strategies in real-world and research environments

COURSE OUTCOMES:

CO 1:	To evaluate scalable AI deployment strategies in real-world and research environments	**Apply**
CO 2:	Analyze containerization, orchestration, and AIOps techniques for intelligent monitoring, automation, and scalable AI system deployment	**Analyze**
CO 3:	Implement optimization techniques in deep learning frameworks for efficient model execution	**Apply**
CO 4:	Analyze distributed AI computing models and performance challenges in large-scale systems	**Analyze**
CO 5:	Analyze real-world AI deployment systems with respect to scalability, performance, and resource utilization	**Analyze**

CO - PO MAPPING:

Particular	PO1	PO2	PO3	PO4	PO5	PO6
CO 1	2	1	3	2	1	3
CO 2	2	2	3	3	1	2
CO 3	2	1	3	3	1	2
CO 4	3	2	3	3	2	2
CO 5	2	2	3	2	2	3

SYLLABUS:

UNIT I: ACCELERATED COMPUTING ARCHITECTURES 9

Parallel computing fundamentals (Amdahl's Law, Gustafson's Law) – GPU architecture: CUDA cores, memory hierarchy, warp scheduling – Mathematical model of parallel speedup and efficiency – Interconnects (PCIe, NVLink, RDMA – analytical view) – AI accelerators (TPU, FPGA – architectural comparison) -- Operating system support for parallel computing – Virtualization overhead analysis – Cloud GPU architecture – Performance bottlenecks in AI systems

UNIT II: AIOPS AND AI SYSTEM ORCHESTRATION 9

Docker architecture (namespaces, cgroups) – GPU containerization (CUDA containers) – Kubernetes architecture (pods, scheduler, control plane) – Resource scheduling algorithms – Auto-scaling strategies in AI systems – MLOps pipelines (CI/CD for ML) – Kubeflow – Monitoring (Prometheus, Grafana – metrics analysis) – Limitations of traditional monitoring – Introduction to AIOps – Anomaly detection & intelligent alerting – Case study: Anomaly detection & intelligent alerting

UNIT III: ACCELERATED DEEP LEARNING AND OPTIMIZATION 9

GPU-based training pipelines – Backpropagation optimization in parallel systems – Mixed precision training (FP16, BF16 – numerical stability) – Memory optimization techniques – Data parallelism vs model parallelism – Distributed gradient descent (mathematical view) – Framework internals (PyTorch vs TensorFlow execution graphs) – Transfer learning optimization – Performance benchmarking metrics

UNIT IV: DISTRIBUTED AI SYSTEMS AND SCALABILITY 9

Distributed computing models (MPI, MapReduce – conceptual comparison) – Multi-GPU and multi-node training architecture – Communication strategies (ring-allreduce, parameter server) – Horovod architecture and working – Scalability analysis (latency, throughput, bandwidth) – Load balancing in distributed AI – Fault tolerance in distributed training – Performance tuning techniques – Case studies (large-scale AI systems)

UNIT V: AI DEPLOYMENT, EDGE SYSTEMS, AND RESEARCH TRENDS 9

Inference optimization (TensorRT – concept) – Edge AI systems and constraints – Federated learning (mathematical formulation) – AI pipelines and ETL optimization – Big data frameworks (Spark, Dask – execution model) – Real-time AI systems (latency constraints) – Smart city and healthcare AI (system-level analysis) – Ethical and sustainable AI deployment – Emerging research trends in accelerated AI

CONTACT PERIODS:

Lecture: 45 Periods **Tutorial:** - Periods **Practical:** - Periods **Project:** - Periods **Total:** 45 Periods

TEXTBOOKS:

1. David B. Kirk and Wen-mei W. Hwu, "Programming Massively Parallel Processors: A Hands-on Approach", 4th Edition, Morgan Kaufmann, 2020
2. Valliappa Lakshmanan, Sara Robinson, and Michael Munn, "Machine Learning Design Patterns", 1st Edition, O'Reilly Media, 2020
3. Jason Sanders and Edward Kandrot, "CUDA by Example: An Introduction to General-Purpose GPU Programming", 1st Edition, Addison-Wesley Professional, 2010

REFERENCES:

1. Tito D'Aquino, J.B. Rainsberger, Nicholas Chase, and Jesse Butler, "Cloud Native MLOps with Kubernetes", 1st Edition, Manning Publications, 2024
2. Roman Trobec, Boštjan Slivnik, Borut Robič, and Patricio Bulić (Eds.), "High Performance Computing for Machine Learning", 1st Edition, Springer, 2021
3. Hollis Denning, "AI GPU Workloads for Beginners: A Practical Guide to Training, Running & Optimizing AI Models on Modern GPUs", 1st Edition, Independently Published, 2025
4. Christopher M. Bishop, "Pattern Recognition and Machine Learning", 1st Edition, Springer, 2006

Head of the Department
Department of Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore - 641 407.

KPR Institute of Engineering and Technology
Centre for Academic Courses
Coimbatore

VERTICAL: Data Sciences

		Category: PEC				
P26DSP05	AI TRiSM	L	T	P	J	C
		3	0	0	0	3

QP PATTERN: B

COURSE OBJECTIVES:

- To develop an understanding of theoretical foundations and challenges in building trustworthy, secure, and reliable AI systems
- To equip students with analytical and evaluative skills for assessing risks, security vulnerabilities, and governance mechanisms in AI
- To enable critical evaluation of real-world AI systems using advanced TRiSM frameworks and research methodologies

COURSE OUTCOMES:

CO 1: Explain theoretical foundations of trustworthy AI, including lifecycle risk management and ethical principles **Understand**

CO 2: Analyze explainability, interpretability, and transparency techniques using formal methods and evaluation metrics **Analyze**

CO 3: Apply statistical and computational methods to assess reliability, uncertainty, and bias in AI models **Apply**

CO 4: Examine security threats and robustness techniques in AI systems using adversarial and probabilistic models **Analyze**

CO 5: Evaluate AI systems and governance frameworks using TRiSM principles, regulatory standards, and real-world case studies **Analyze**

CO - PO MAPPING:

Particular	PO1	PO2	PO3	PO4	PO5	PO6
CO 1	2	1	3	2	1	3
CO 2	2	2	3	3	1	2
CO 3	2	1	3	3	1	2
CO 4	3	2	3	3	2	2
CO 5	2	2	3	2	3	3

SYLLABUS:

UNIT I: THEORETICAL FOUNDATIONS OF TRUSTWORTHY AI AND TRISM 9

Concepts of Trustworthy AI (fairness, accountability, robustness) – AI lifecycle and TRiSM framework – Risk modeling basics (Risk = Probability × Impact) – Introduction to uncertainty and risk in AI systems – Ethical principles in AI (utilitarianism, deontology) – Stakeholders and accountability in AI ecosystems – Overview of AI regulations (NIST AI RMF, EU AI Act) – Challenges in developing trustworthy AI systems – Case discussions and emerging trends in trustworthy AI

Learn Beyond

UNIT II: EXPLAINABILITY, INTERPRETABILITY, AND MODEL TRANSPARENCY 9

Need for explainable AI (XAI) in modern systems – Black-box vs interpretable models – LIME: local interpretability approach (concept and workflow) – SHAP: feature contribution analysis (intuitive understanding) – Grad-CAM and visualization techniques in deep learning – Model-agnostic vs model-specific explanation methods – Evaluation metrics for explanations (fidelity, consistency, stability) – Trade-off between accuracy and interpretability – Limitations of XAI and research paper discussion

UNIT III: RELIABILITY, UNCERTAINTY, AND BIAS IN AI SYSTEMS 9

Bias-variance trade-off in machine learning – Types of bias in AI systems (data bias, algorithmic bias) – Feature selection techniques (correlation-based methods) – Handling heterogeneous and noisy data – Model evaluation metrics (accuracy, precision, recall, F1-score) – Validation techniques (cross-validation, train-test split) – Introduction to uncertainty in AI (epistemic vs aleatoric) – Reliability and robustness in AI models – Case studies on model failure and generalization issues

UNIT IV: AI SECURITY, ADVERSARIAL LEARNING, AND ROBUSTNESS 9

Overview of AI threat landscape – Adversarial attacks (FGSM – conceptual understanding) – Data poisoning and backdoor attacks – Model extraction and inference attacks – Risk identification and classification in AI systems – Risk mitigation strategies and defense mechanisms – Robustness techniques in AI models (high-level approaches) – Security evaluation metrics and testing strategies – Case studies and research insights on AI security failures

UNIT V: AI GOVERNANCE, RISK EVALUATION, AND COMPLIANCE FRAMEWORKS 9

Fundamentals of AI governance and policy frameworks – NIST AI Risk Management Framework (structured understanding) – Regulatory frameworks (GDPR, EU AI Act – overview) – Compliance and legal considerations in AI deployment – AI auditing and accountability mechanisms – Risk assessment and impact analysis models (conceptual) – Ethical decision-making in AI systems – Trade-offs in AI systems (performance vs fairness vs compliance) – Case study evaluation and future trends in AI governance

CONTACT PERIODS:

Lecture: 45 Periods **Tutorial:** - Periods **Practical:** - Periods **Project:** - Periods **Total:** 45 Periods

TEXTBOOKS:

1. R. Karthick Manoj, S. Senthilnathan, S. Arunmozhi Selvi, T. Ananth Kumar, "AI Trust, Risk, and Security Management", Wiley-Scrivener, 2024

REFERENCES:

1. Qinghua Lu, Liming Zhu, "Responsible AI: Best Practices for Creating Trustworthy AI Systems", Pearson, 2023
2. Kate Crawford, "Atlas of AI", Yale University Press, 2021
3. Arvind Narayanan, Sayash Kapoor, "AI Snake Oil", Princeton University Press, 2024
4. Huan Liu et al., "Trustworthy Machine Learning", Morgan Kaufmann, 2022

Head of the Department
Department of Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore - 641 407

KPR Institute of Engineering and Technology
Centre for Academic Courses
Coimbatore

Learn Beyond

VERTICAL: Data Sciences

		Category: PEC				
P26DSP06	NoSQL Database	L	T	P	J	C
		3	0	0	0	3

QP PATTERN: B

COURSE OBJECTIVES:

- Understand limitations of relational databases and need for NoSQL
- Learn different NoSQL data models and architectures
- Apply NoSQL concepts using MongoDB
- Design scalable and distributed database solutions

COURSE OUTCOMES:

CO 1: Explain NoSQL concepts and data models given fundamental database scenarios and examples with clarity and correctness **Understand**

CO 2: Analyze distributed architectures and consistency models given different NoSQL system configurations with appropriate reasoning and interpretation **Analyze**

CO 3: Apply MongoDB operations (CRUD, Queries, Indexing) using MongoDB tools and datasets to produce correct and efficient outputs **Apply**

CO 4: Examine and model scalable NoSQL database solutions based on given application requirements with appropriate design considerations **Analyze**

CO 5: Analyze and compare different NoSQL systems for suitability in real-world applications with justified conclusions **Analyze**

CO - PO MAPPING:

Particular	PO1	PO2	PO3	PO4	PO5	PO6
CO 1	1	-	3	2	-	-
CO 2	2	-	3	3	-	1
CO 3	2	-	3	3	-	-
CO 4	3	-	3	3	1	1
CO 5	2	1	3	3	2	2

SYLLABUS:

UNIT I: INTRODUCTION TO NOSQL DATABASES **9**

Evolution from RDBMS to NoSQL – Characteristics of Big Data and Distributed Computing – Need of NoSQL databases – Characterisitics of NoSQL databases – Types of NoSQL databases: Key-value, Document, Column-family, Graph – CAP Theorem – BASE properties – Comparison between SQL and NoSQL Databases – Overview of NoSQL Architecture – Applications of NoSQL Databases

UNIT II: KEY-VALUE DATABASES 9

Introduction to Key-Value Stores – Architecture and Working Principles – Data Storage and Retrieval – Hash Tables and Distributed Hashing – Replication and Partitioning – Consistency Models – Operations in Key-Value Databases – Use Cases and Applications – Introduction to Redis and Riak – Basic CRUD Operations

UNIT III: DOCUMENT DATABASES & MONGODB BASICS 9

Introduction to Document databases – Introduction to MongoDB architecture – JSON and BSON format and collections – CRUD operations in MongoDB – Querying and projections – Indexing in MongoDB – Aggregation Frameworks – Replication and Sharding – Applications of Document Databases

UNIT IV: COLUMN-FAMILY DATABASES 9

Introduction to Column-Oriented Databases – Wide Column Store Concepts – Column Family Architecture – Distributed Storage Mechanisms – Data Modeling Techniques – Consistency and Scalability – Introduction to Cassandra and HBase – CQL (Cassandra Query Language) – Data Partitioning and Replication – Applications of Column Databases

UNIT V: SPECIALIZED NOSQL SYSTEMS & APPLICATIONS 9

Introduction to Graph Databases – Nodes, Edges, and Properties – Graph Data Models – Graph Traversal Techniques – Query Languages for Graph Databases – Introduction to Neo4j – Cypher Query Language – Social Network and Recommendation Applications – Graph Analytics and Real-Time Processing – Advantages and Challenges of Graph Databases

CONTACT PERIODS:

Lecture: 45 Periods **Tutorial:** - Periods **Practical:** - Periods **Project:** - Periods **Total:** 45 Periods

TEXTBOOKS:

1. Michael Kaufmann, Andreas Meier, "SQL and NoSQL Databases : Modeling, Languages, Security and Architectures for Big Data Manage", 2nd Edition, Springer Nature Switzerland AG (Springer, Cham), 2023
2. Sriram, Dhanusha C., Kanagalakshmi S., "Mastering NoSQL with MongoDB", 1st edition, Deep Science publishing, 2025

REFERENCES:

1. Robert Johnson, "SQL and NoSQL: Building Hybrid Data Solutions for Modern Applications", 1st edition, HiTeX Press, 2024
2. Greyson. Chesterfield, "Mastering NoSQL Databases with MongoDB: Building Flexible, Scalable, and High-Performance Databases", 1st edition, Amazon Digital Services LLC - Kdp, 2024

Head of the Department
Department of Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore - 641 407.


Learn Beyond